Inferring Human Dynamics in Slums Using Mobile Phone Data

Amy P. Wesolowski

College of the Atlantic, 105 Eden St., Bar Harbor, ME 04069 and

Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501

Nathan Eagle

Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501

Over one billion people live in the world's 200,000 slums and informal settlements. We use mobile phone data to better understand this growing population. In particular, we consider Kibera, located in Nairobi, Kenya. Using mobile phone call logs from June 2008 to June 2009 and theories by human geographers, economists, sociologists, journalists, and anthropologists as a basis, we tested the validity of a few prominent theories. In particular, we focus our research on inferring places of work and migration out of Kibera.

I. INTRODUCTION

How much can a mobile phone tell us about a person? a group of people? a society? a culture? a nation? In America alone, the impact of mobile phones is clear. They have become an ingrained part of our culture, practically necessary to function in our society. Mobile phones have gone from a luxury to an extension of nearly every person in our country. Currently, there are over four billion mobile phone subscribers around the world. However, the majority of these are in the developing world, where the rate of adoption greatly outpaces that of the developed world. In 2008, the number of mobile phone subscribers in Africa passed the number in North America with over 280 million subscribers.

As a result, mobile phones are providing enormous behavioral data sets, especially for underrepresented populations in the developing world. Using mobile phone call logs, we can track human movement, infer socioeconomic status, and better understand human dynamics. As opposed to self-reported surveys or anecdotal evidence from field work, human behavior can be quantified using these data sets without human bias. Recently, work has been done to use mobile phones to increase our insight into human movement and behavior[7] [3].

This research focuses on mobile phone data generated in Kenya from June 2008 to June 2009. Currently, Kenya is the fastest growing mobile phone market in the world. From 1999 to 2006, the number of mobile phones grew from 15,000 to 5.6 million. The adoption rate in Kenya is allowing us to look at severely underrepresented populations. In particular, we are focusing on the lives of Kibera's, the largest slum in Nairboi, Kenya, residents. Our research is working towards building a better understanding of migration out of Kibera and places of work. We are using call patterns to better understand these two driving forces that will enable us to quantify Kibera's dynamics. Ultimately, we plan on using aggregate statistics found in this paper to motivate mathematical models and develop measures to quantify slum dynamics.

A. Why Study Slums?

> For the first time in history, more people live in cities than in the countryside. Our world is no longer simply going through the experience of urbanization. Our world has become urbanized. ... One billion people - or one in every three urban residents - now live in an urban slum, the vast majority of them in developing nations[10].

By 2015, there will be at least 500 cities whose population is over one million[16]. It is estimated that by 2050, the world population will reach ten billion, with the majority of those people living in urban areas[5]. The brunt of this population growth will occur in developing countries. Ninety-five percent of the growth of the human population will occur in the urban areas of developing countries, whose population is expected to double to nearly four billion over the next generation[20].

However, anthropologists and human geographers agree that we have limited knowledge about what effect this growth will have on the world. Unlike the growth of cities during the industrial revolutions in Europe and North America, as a population we have never experienced such profound growth. Social scientists are working on better understanding this growth and change, however little quantitative work has been done. We believe that this work is step in the right direction towards quantitative results measuring slum dynamics.

The United Nations Human Settlements Programme (UN-HABITAT), has tried to better understand and address this mounting concern. In particular they have focused their research on the growth of slums, poverty, and the effect these two have on developing countries. Although the definition of a slum varies from country to country, we are using a generally accepted definition

based on prominent urban theories. A squatter settlement, or slum, can be defined as "a residential area which has developed without legal claims to the land and/or permission from the concerned authorities to build; as a result of their illegal or semi-legal status, infrastructure and services are usually inadequate[6]". It is clear that the growth and proliferation of cities has coincided with the emergence of larger and denser slums.

Currently, there are over one billion slum residents worldwide. However, little research has been about this increasing population. In *The Challenge of Slums*, the UN employed more than a hundred researches conducting case-studies of poverty, slum conditions, and housing policy. They produced governmental reports citing the importance of understanding slums, poverty, and the effect it has on urban areas. Their work has been done on a much larger scale, in terms of hundreds of people surveyed, then previous social scientsts. However, it is extremely difficult to capture a representative sample size, even if you only consider a single country.

Since, urban areas, in particular slums, provide a means for individuals to improve their quality of their life, their growth is inevitable. With the growth of slums comes an array of issues, mostly concerning health and safety concerns for slum residents. There are a number of physical, social, and legal characteristics that define a squatter settlement including a lack of social infrastructure, public facilities, lower income earning residents, and an illegal or informal housing agreement. With a lack of governmental infrastructure and plans for urban growth, it is unsurprising that slums continue to prosper. In most cities, government officials view slums as a problem that is largely ignored. Little work is done to understand, serve, and control the growth of slums. The problem is further compounded by the apathy and even anti-pathy of various government agencies who view the "invasion" of urban areas by "the masses" and the development of squatter settlements as a social "evil" that has to be "eradicated".

In Nairboi alone, nearly 2 million people live in its informal settlements and slums. Representing around half of the total population of Nairobi, these individuals live in 66 uniquely identified slums. In Kibera, the largest slum, population estimates vary, with most agreeing the population is well over 600,000 or around one-fifth of Nairobi's population[10]. These residents live in only 5 percent of the total residential area without basic necessities, such as sanitation, water, healthcare, or education[1]. However, even with these conditions, the population of Kibera grows at an annual rate of 12 percent[10]. However, the Kenyan government excludes slums from city authority planning and budgeting and treats these regions as if the do not exist.

Our main research goals are to be able to quantitatively infer places of work and migration patterns out of Kibera. Both of these topics are key to building a model that accurately reflects slums. In particular, understanding migration dynamics in informal settlements

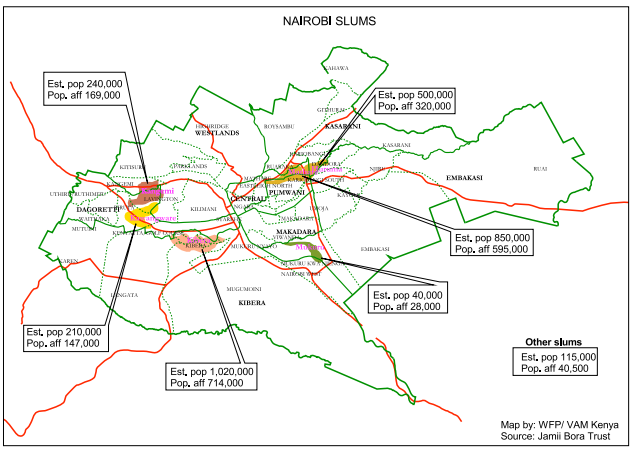

FIG. 1: This is a map of Nairobi, Kenya with the largest slums and estimated populations marked.

is of great importance. One of the African Population and Heath Research Center's research programs; Urbanization, Poverty, and Health Dynamics in sub-Saharan Africa (UPHD) lists an understanding of migration dynamics in Nairobi as a key piece to better understanding the health consequences of rapid urbanization. They are working to determine the characteristics of in- and out- migrants, motivations for migration, connections between their origin communities, and how poverty is influenced by intra-urban and urban-rural migrations. Their data is a combination of government statistics, NGO field reports, demographic surveillance systems, and self-reported surveys. Our research is attempting to address some of the same questions using mobile phone data.

However, the characteristics of slums and typical slum residents make movement patterns difficult to quantify. From temporary work situations, unstable tenure housing, and a lifestyle driven by available work, slum residents live extremely transient lives. Nonetheless, we have found overarching trends that can further classify and understand slum movement.

II. DIFFERENT MODELS OF MIGRATION

Our research works to quantify human mobility in slums. We are using many theories from human geography, economics, anthropology, and urban studies to motivate our questions. Below is an overview of the prominent theories of migration based on field work, government statistics, and NGO documents.

In Goodman's review of McGee's work, *The Urbanization Process in the Third World: Explorations in Search of a Theory*, "of all that social scientists have said about urbanization in developing countries, few would disagree that most is hearsay. The welter of propositions and hypotheses, built on irregular censuses and sketchy data banks, far exceed the ability of such data to adequately

support them[8]." Through our work, we are moving to better quantify and find evidence to support these theories. First we will review some ideas about an individual's motivation for migration and then make broader statements about the general population.

If we consider the migration of an individual, much work has been done on studying an individual's motivation for migration.

> Although seen to flow from less to more advantageous places, migration is usually studied as independent of links with development or ubanisation. Guiding factors include wage and job opportunity differentials in the modern sector; employment opportunities in the informal and rural nonfarm or small scale enterprise sectors; migration chains based upon family, extended family, and acquaintance relationships; circular and seasonal migration strategies; individual or household characteristics such as age and family size; and resource-push factors related to the origin town or village: its economic well-being, pattern of resource distribution among social classes, local norm, and its integration into the urban network[2].

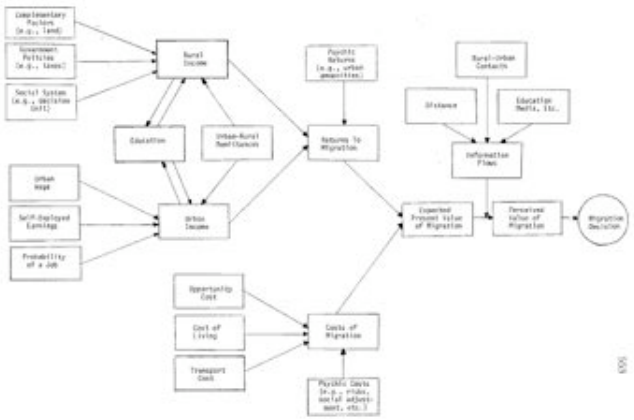

FIG. 2: Migration decision chart from [4]

Work appears to be the key motivation for migration. However, working environment must go hand in hand with living environments. Most current theories are built on Turner's original theory of urban slum migration. Turner proposes that "there are three basic functions of the dwelling environment: *location, tenure* and *amenity*. For any place to function as a dwelling it must have an accessible location, it must provide secure continued residence for a minimum period, and it must provide a minimum of shelter from hostile elements – whether climatic or social[15]." From these functions, he then identifies three stages in the trajectory of the urban poor: bridge-header, consolidator, status-seeker. The poor in each of these stages prioritize their needs and this greatly affects their migration patterns. For example, the bridge-headers are most concerned with income-generating opportunities, whereas the status-seekers are most concerned with modern amenities as opposed to location or tenure.

Overall, the theories can be summarized as follows:

1. Theory: people move from rural areas to urban slums, work and then move to the outskirts of a city. This theory relies on the idea that when people move into slums, they are able to increase their socioeconomic status.

2. Theory: people move from rural areas to urban slums, work and then move to other slums. This theory views a slum as a sink, as opposed to a filter. The general idea is that when people move into slums, they are unable to increase their status and are forced to continue to move around other slums.

3. Theory: people move from rural areas to urban slums and stay in the slums. This theory views a slum as the final destination. Here individuals are more like to remain in a slum once they move there for a number of economic, social, and environmental reasons.

4. Theory: people move from rural areas to urban slums and then back to rural areas. This is the popular theory of urban-rural migration where there is a large proportion of the slum population that continually moves back and forth from rural villages.

A. Rural-Urban Migration

By far the most popular theory of African migration, rural-urban migration has a rich history that has greatly shaped the current economic, political, and social structure of many African countries. "In Africa, the migration of labor both in colonial times and presently has been particularly important in the development process[4]." Historically, younger men have migrated from their villages for a period of several years to work in urban areas or on plantations. However, that trend has since changed as unemployment has increased in urban areas. Now, it is more common for individuals to be permanent migrants who rarely travel back to their original villages. Although most agree with the impact of rural-urban migration, most theorists differ in their model. Traditionally, the models closely resemble a cost-return analysis where the cost of migration is weighted against the return of income. Byerlee's model differs from a conventional cost-return analysis by including elements of the social system, identifying determinants of rural and urban incomes, and introducing risk and costs of migration. Byerlee weighs each of these factors to make a more realistic representation of an individual's motivation and the effect it has on migration. Even though employment

is the prime motivation for rural-urban migration, it has resulted in excess labor. "On the contrary, migration today remains a major factor contributing to the phenomenon of urban surplus labor; a force that continues to exacerbate already serious urban unemployment problems caused by the growing economic and structural imbalances between African urban and rural areas[14]."

Most cite rural-urban migration as a leading cause for slums, the growth of cities, and agglomeration. Van der Ploeg and Poelhekke study the proliferation of megacities. They state that the largest portion of "city growth is due to the rapid growth in informal employment and informal housing (slums). People move to a city in the hope of finding work and improving their real wage ... Their benchmark is the expected real wage they can obtain outside the city, either in the rural sector if most of the country is still rural or in another city's wage if urbanization has already reached high levels[21]." This motivation has caused cities to become their current size and continue to grow.

Agreeing with Byerlee, Todaro cites population growth and an accelerated rural-to-urban migration as the chief causes for the expansion of slums. "... the informal sector (slums) is linked with the rural sector in that it allows unskilled laborers to escape from rural poverty and underemployment, although it grants them living and working conditions and incomes that are not much better than what they had had before moving. The question remains, however, as to whether the informal sector is merely a holding ground for people awaiting entry into the formal sector (middle class areas) and, as such, is a transitional phase that must be made as comfortable as possible until it is absorbed by the formal sector, or whether it is here to stay and should, in fact, be promoted as a major source of employment and income for the urban labor force."

Aside from employment, rural-urban migration is greatly affected by tribal affiliations. It is widely accepted that people are more likely to move to a slum where they know others from their tribe already reside. This theory makes intuitive sense. Silveria et al. analyze this dimension of migration using an agent-based model. As opposed to previous models, the authors not only take into account economic incentives, they also take into account the effect of social neighborhoods[13]. However, little work has been done to quantify the strength of this relationship or assess its impact. Our research attempts to determine to what extent tribal affiliations effects the dynamics of slums.

B. Urban-Urban Migration

As opposed to rural-urban migration, Brown and Moore have studied the intra-urban migration processes. Their focus is on behavioral decision making processes of individuals that result in migration. They highlight factors necessary to better understand an individual's decision to move within an urban setting. The first level of decision making should be comprised of measures looking at the relationship between household characteristics and needs (such as wealth, socio-economic status, family life), stressors encouraging movement (such as political pressures, time constraints), and spatial expressions (i.e. the relationship and reaction to a given neighborhood). After this step of assessment, site determination is comprised of accessibility, physical characteristics, availability of services and facilities, the social environment, and individual site and dwelling characteristics. Moreover, they propose an aspiration level associated with intra-urban migration. It is assumed that contributing factors are an increase in social status, the probability of finding a vacancy, and the sense of community that can be found in a new dwelling. They also highlight the utilization of past information, the probability of success, amount of perceived effort, and the amount of time individuals have to make a decision. Overall, they encourage a division of individuals based on social, economic, and locational dimensions that greatly affect their information channel and level of awareness[2].

Coupling Brown, Moore, and Turner's theories, when individuals move out of Kibera and stay in the urban settling, they should be moving to residential areas that provide more stable housing, work opportunities, and a higher quality of living. However, studies have not been conducted to the provide evidence for this theory. Moreover, none of the prominent theorists have provided estimates on the ratio of rural-urban to urban-urban migration. Our work attempts to address this fact.

III. REGIONAL AFFILIATIONS

In order to infer tribal affiliations in our analysis, we are associating tribes with regional dialects. Although English and Swahili are the official languages in Kenya, there are currently thirty distinct languages/dialects spoken in Kenya[1]. These are divided into four main linguistic groups: Bantu, Nilotic, Cushitic, and Swahili. The linguistic divisions of Kenya correlate to geographic locations and are shown in Figure 3.

IV. THE DATA

We are looking at call logs made in Kenya from June 2008 to June 2009 on a monthly basis[2]. Our data is a log of each call made in Kenya with the person calling, person called, tower the person called from, date, and duration[3]. In total we have around ten billion calls logs representing over 6 million mobile phone subscribers.

From these files we created mobility files. Mobility files are comprised of a caller, tower ID, region associated with the tower, and date of the call. We then consider the mobility files created for each month.

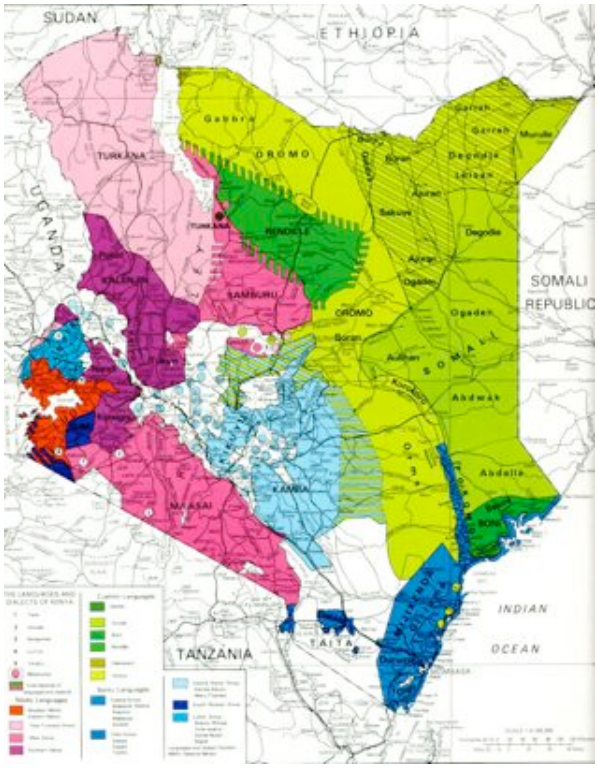

FIG. 3: This is a map of Kenya divided by dialect.

TABLE I: Number of towers for each region.

Region	Number of Towers
Boni	47
Mijikenda	856
Pokomo	12
Oromoi	6
Rendille	3
Somali/Arjuran	3
Samburu	3
Turkan	3
Kalenjin	6
Kayia	111
P Next to Kayia	11
Luo	100
Guss	282
Kipsigis	353
Nandi	5
Maaasi	1327
Kamba	3
Nairobi	4743
Unknown Region[a]	3123

[a]Unknown Region represents towers located in areas without a prominent tribal affiliation.

All of these calls are registered at one of over eleven thousand call towers in Kenya. The majority of these are in the southern and coastal areas of the country where the population is the densest. These unique eleven thousand cell towers are in 2,700 locations. We have correlated cell towers with regions and tribal affiliations.

V. KIBERA

Our research focuses on a particular slum in Nairobi, Kenya. Kibera is the second largest slum in Africa located southwest of Nairobi's city centre. Founded around 1912 as a gift to Nubian soldiers by the British military, this land is considered a forest by the Kenyan government. There are currently between 400,000 and one million residents living in an area roughly 75% that of Manhattan's Central Park (about 550 acres). Over 90% of residents are tenants living in Kibera's many villages.

Numerous sociologists, anthropologists, and journalists have studied Kibera. Most rely on anecdotal evidence and personal reflections on the slum to make assessments about its dynamics. Also, the majority of research has been conducted with the aim of helping residents secure tenure housing, clean water, and proper sanitation. This work is valuable and helps guide our research questions. However, quantitative work enables us to make statistically significant statements about the changes in the slum on a much larger scale.

We are working on two main topics: migration out of Kibera and inferring places of work. It is interesting to study Kibera because it has remained surprisingly consistent over the last ninety years. Neuwirth states that "Of all the shadow cities, I visited, the shantytowns of Kenya are the only ones that are not really squatter cities. And they are the only ones that have remained stagnant, the living conditions largely uncharged for half a century[11]." Even though it is well accepted that migration plays a major role in the growth of Kibera, the slum has existed in a consistent state for many years. We want to know what affect these two key factors have on Kibera. Is there little movement within in the slum? Is there little migration in and out of Kibera? These are questions we will answer and ultimately use these results as motivation for future work.

Our first goal is to determine individuals living in Kibera. To do this we use the only cell tower location inside of Kibera. This cell tower location, 307, has six unique cell tower IDs associated with it. We will refer to this tower as T_k.

VI. CLASSIFICATIONS OF CALLERS

We classify a caller as living in Kibera if they meet all of the following criteria:

1. Over fifty percent of their total calls between the hours of 6 PM and 8 AM have been made at T_k.

FIG. 4: Google Earth image of Kibera with cell phone tower site IDs marked

2. The total number of calls made in a month is between 3 and two standard deviations from the mean number of calls made by those living in Kibera[4].

By our classification scheme we are looking at a sample set of around 18,000 total callers in Kibera over the course of the year[5].

As Figures 5 and 6 show, there are between 2090 and 3068 callers in Kibera making between 53,868 and 74,489 calls per month.

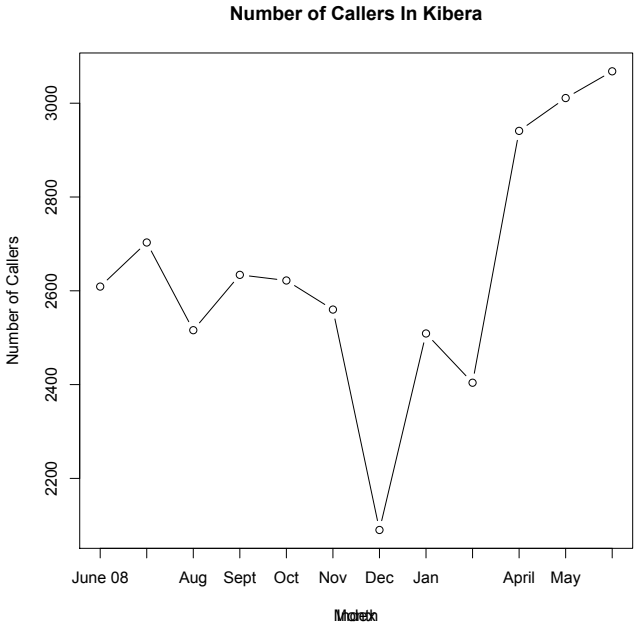

FIG. 5: This graph represents the total number of callers per month living in Kibera.

We are further classifying callers in Kibera by the following:

1. Let $C_1 = \{x : x \in M_1, x \in M_2\}$ where M_1, M_2 represent consecutive months. C_1 is the set of individuals who are living in the slum from one month to the next month.

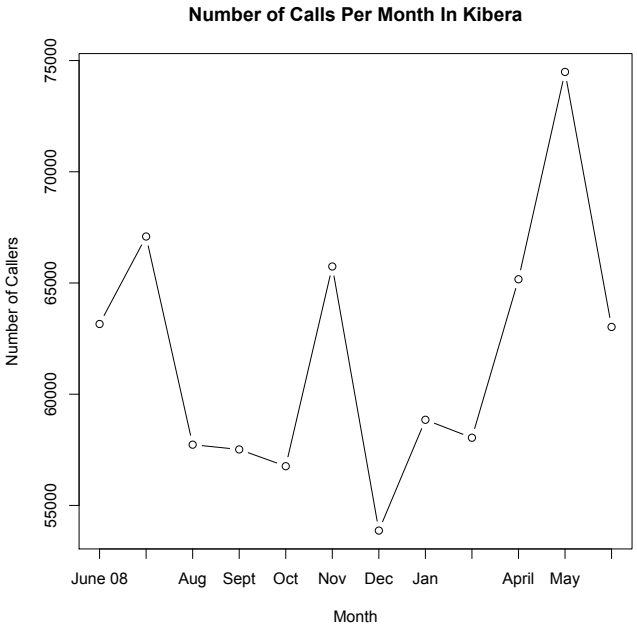

FIG. 6: This graph represents the total number of calls made per month by individuals living in Kibera.

2. Let $C_2 = \{x : x \notin M_1, x \in M_2\}$. C_2 is the set of individuals who are not in M_1, but are in M_2, i.e. those who have moved into the slum during M_2.

3. Let $C_3 = \{x : x \in M_1, x \notin M_2\}$. C_3 is the set of individuals who are in the first month M_1, but not in the second month M_2, i.e. those who have moved from the slum before the start of M_2[6].

As Figure 8 shows, the majority of people live in Kibera for less than two months, with the mean at 1.950 months. This result supports the theory that slums are more similiar to a filter as opposed to a sink. This is evidence that Kibera has an extremely high turn over rate month to month. We refine this figure later in the paper.

For all callers in C_3, we want to determine their new residence. To classify a person's living region, we consider an individuals entire call history for a given month. We create regions by towers that are within 1 km of one another and then sum the total number of calls per region. This will group towers in close proximity to one another. We are considering a person's most frequently called from region as their residence region. For much of our analysis, we want to determine overlapping regions or regions where the majority of Kibera's residents are calling from. In order to accomplish this, we take the intersection of a set of callers' regions.

Furthermore, in order to classify individuals who have moved from the center of Kibera to another section of Kibera, we have added all towers on the border of the slum into the slum tower list. We refer to this set of towers as T_e. We then further divide C_2, C_3 into two

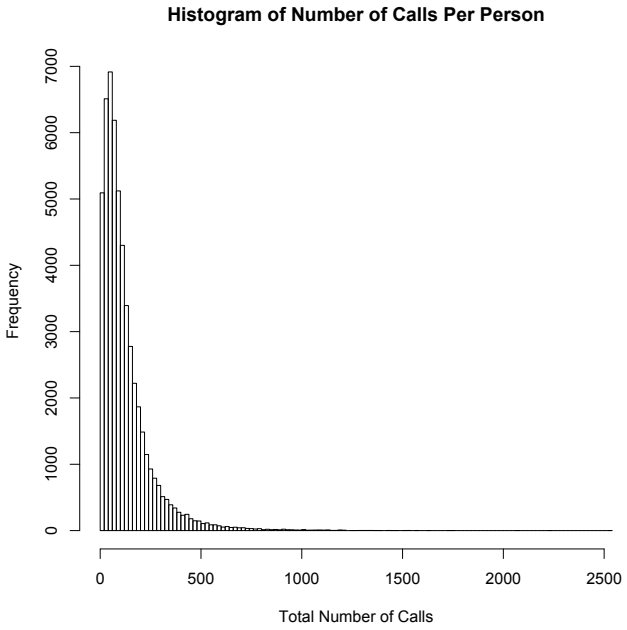

FIG. 7: This graph represents the number of calls made for each caller living in Kibera in May 2009. Overall, the mean for each caller is around 25 calls per month

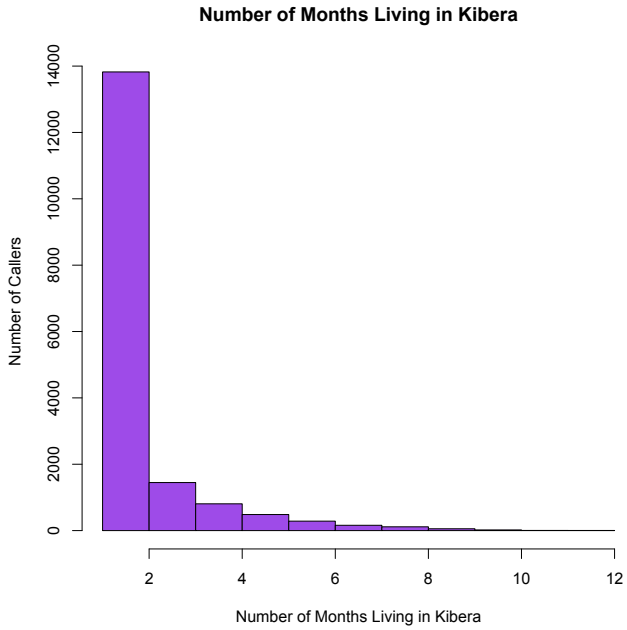

FIG. 8: This chart represents the number of people who have lived in Kibera for a specific number of total months.

subsets each.

1. Let $c_f \subseteq C_2$. c_f is the set of people whose previous living region does not include any tower in T_e.

2. Let $c_n \subseteq C_2$. c_n is the set of people whose previous living region includes towers in $T_e \backslash T_k$.

3. Let $c_g \subseteq C_3$. c_g is the set of people whose living region does not include any tower in T_e.

4. Let $c_s \subseteq C_3$. c_s is the set of people whose living region includes towers in $T_e \backslash T_k$.

VII. TIME OF DAY

Before considering where people are moving and working, we wanted to determine the distribution of calls made in T_e and elsewhere. When first comparing the number of calls made during the day (i.e. between 8 am and 6 pm, exclusive) and those calls made during the night (i.e. between 6 pm and 8 am, inclusive), the average ratio is 56.375% to 43.625%. As Figure 9 shows, the overwhelming majority of calls made during the day are to other towers in T_e. This fact is important when considering where people are working.

VIII. TRIBAL AFFILIATIONS

Similar to the rest of Kenya, Kibera is greatly effected by tribal affiliations. Its historical roots put its ties with the original homeland of Nubian soldiers. "Even today, 90 years on, several of Kibera's oldest villages ... are dominated by descendants of the original riflemen[11]." However, the majority of residents are not Nubian, but a mixture of tribes. "Urban migration has made Kibera more ethnically heterogeneous, and, amid poverty and the increasing demand for housing and land, ethnic tensions have grown[10]." These tribal disputes are a main cause of violence in the slum. Historically, the Nubian descendants and Luo, the second largest presence in the community originally from the Lake Victorian region, have caused numerous clashes[11]. In late November and early December of 2001, tribal disputes caused massive looting and robbing, causing the local police to go on their own rampage, raping, beating, and destroying property. As a result thousands fled Kibera. With the huge impact of tribal affiliations in Kibera, a better understanding of relationships between tribes and residents is necessary.

In order to infer tribal affiliations in Kibera, we are considering callers who have moved away from Kibera, c_g. We then calculate their new residence region and associate each region with a tribal dialect. First, let us consider the distribution of all regions.

As we can see, the majority of residents have moved to Nairobi or an unknown region of Kenya. Later in the paper, we will take a closer look at movement within

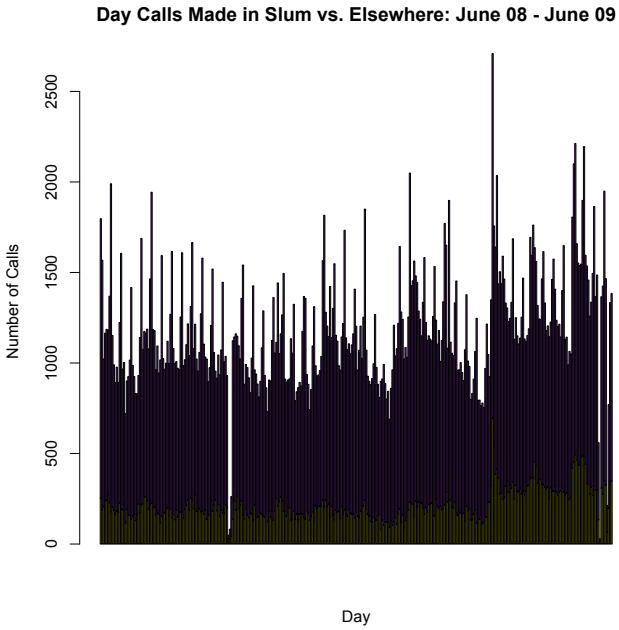

FIG. 9: This graph shows the distribution of calls made during the day in the slum versus elsewhere. The purple portion of the graph represents the number of calls made from T_e. The yellow portion of the graph represents the number of calls made to a tower outside of the slum.

Nairobi. Now, if we remove these top regions (and towers called at unknown locations[7]), we can see that the Maasai are the overwhelming majority. From this graph, we cannot see a clear influence of the Luo. However, it is interesting to note that the regions represented for the entire year remains constant.

Moreover, we have observed that the majority of tribal regions where individuals have moved to are in close proximity to Nairobi. We are further refining our assessment of tribal affiliations to test if people are more likely to move to close regions or regions with a strong tribal pull.

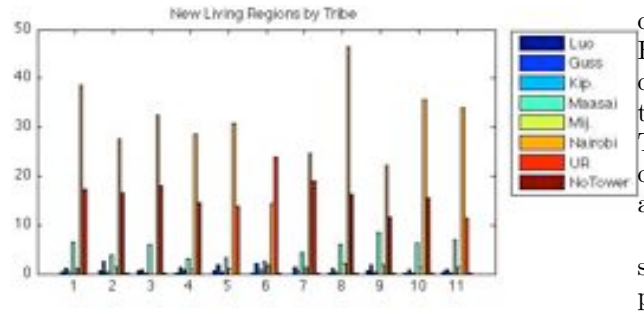

FIG. 10: This graph represents the percentage of residence region calls associated with each dialect division.

This is one method to infer tribal affiliations of Kibera. However, we realize there are flaws in this system. We are not considering calls made to these different regions from Kibera. We are currently working on associating a tribe with each person. We will then take a closer look at calls made from Kibera's residents and inferring tribal affiliations of individuals Kibera's residents have called. We believe this will provide a better assessment of tribal affiliations of Kibera.

IX. URBAN MIGRATION

Before attempting to infer what types of living establishments Kibera's residents are moving to, we want to see the number of times a person moves from Kibera.

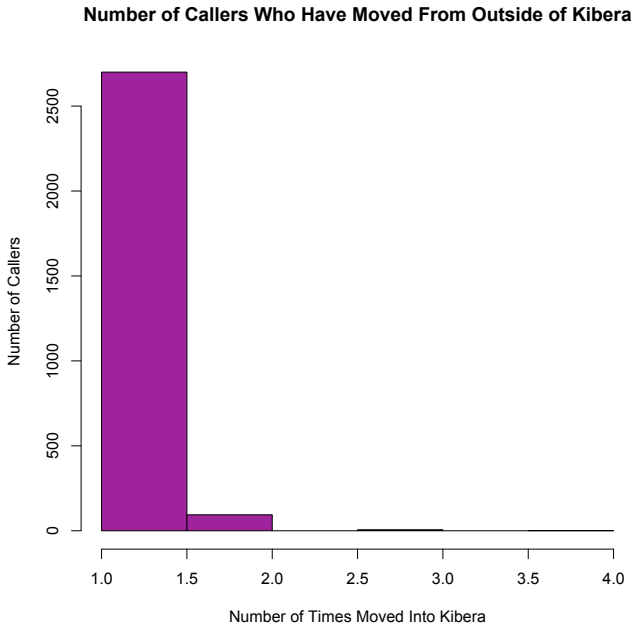

FIG. 11: Number of times a person has moved into Kibera from outside of the slum.

As Figures 11 and 12 show, very few people move from outside of Kibera or to outside of Kibera more than once. Here the mean number of times a person moves from outside of Kibera is 1.039. For the individuals who move to outside of Kibera, the mean number of times is 1.048. This result shows that the majority of people who move out of Kibera do not move back into Kibera and out again.

As Figure 13 shows the mean number of months a person is 'living' in Kibera is now 1.559 and there are more people living for longer amounts of time. However, these results still show that the average time 'living' in Kibera is still very low and there is further evidence of a high turn over rate in the slum.

Many theorists write about the back and forth nature of slum residents – i.e. people move from rural areas to urban areas often. We wanted to see how many times a person is likely to move from living in Kibera to either liv-

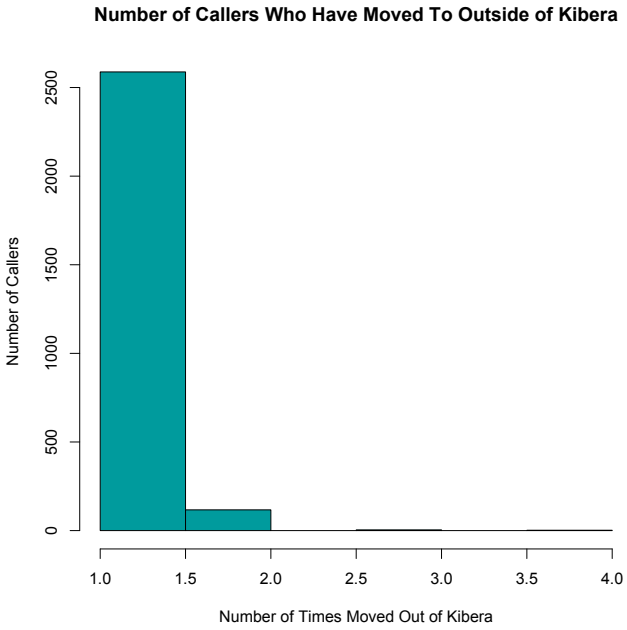

FIG. 12: Number of times a person has moved from Kibera to outside of the slum.

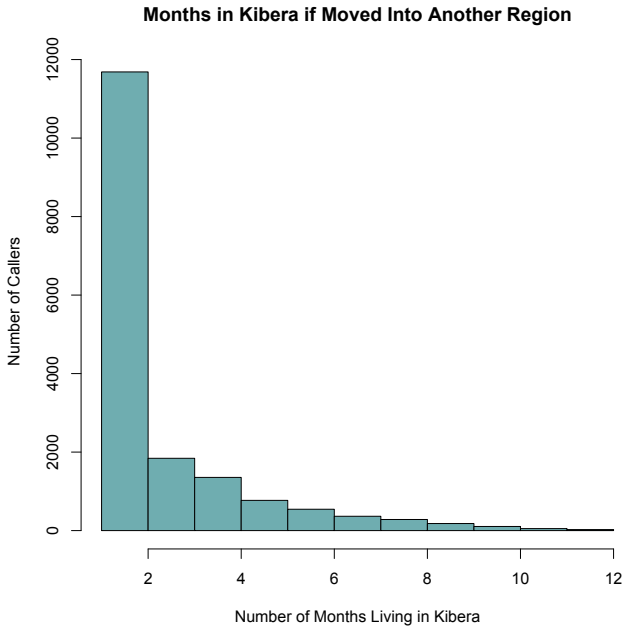

FIG. 13: If we consider the people who have moved out of Kibera, but stayed in another part of Kibera, and the number of people who live in Kibera for a month, this is the number of months a person is 'living' in Kibera.

ing in another part of Kibera or outside of Kibera. These graphs show that the majority of people who move out of Kibera, only do so once. Since we are only considering one years worth of data, any seasonal migration would require a longer longitudinal data set.

As we can see in Figure 14, those who move out of Kibera are not moving a great distance. However, when we take a closer look at these new places of residence, as in Figures 14, we see that the majority of these establishments do not appear to be places of residence. This raises two interesting points, but not the points we intended to make. Firstly, this shows that Kibera's residents maintain many of the same ties they had while living in the slum. The overlap in these regions and the top regions where people are working is an indication of this fact. Moreover, since these new places of residence are close to Kibera, we suspect that many of these people maintain their previous jobs, even if they have moved from the slum. We are currently working on new methods to test this theory. The other interesting point involves our assumption about a person's living region. We assumed that a person's top region represented where they lived. At first, this assumption seemed intuitive. However, it seems as though this may not be an accurate assumption. We are seeing evidence that slum residents' top regions are more likely places of work. Therefore, we must refine our method of inferring living regions.

FIG. 14: This Google image is of the top new living regions in Nairobi of those who have moved out of Kibera.

A. Migration Within a Slum

As Figures 15 and 16 show, the majority of individuals who have moved out of Kibera, have actually moved to other regions of Kibera. This shows that while there is a large proportion of people moving from our original definition of living in Kibera, the majority of these people just add to changes in different areas of Kibera.

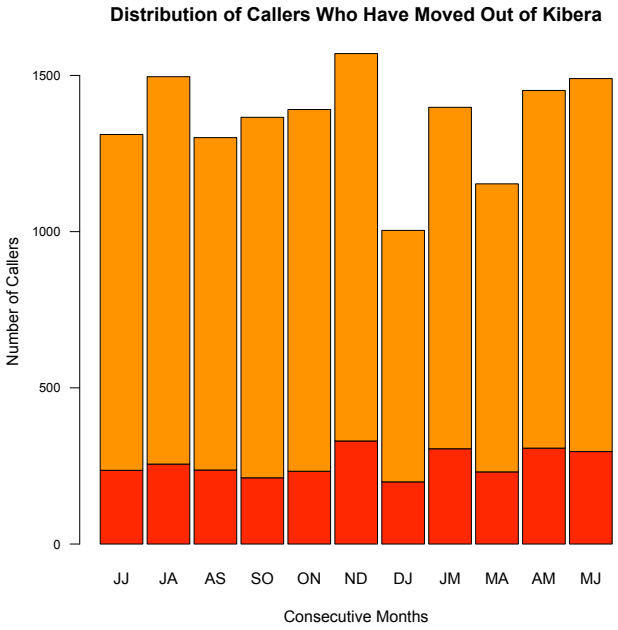

FIG. 15: Distribution of calls who have moved out of Kibera. The orange bar represents those who have moved to another region of Kibera. The red bar represents those who have completely moved out of Kibera.

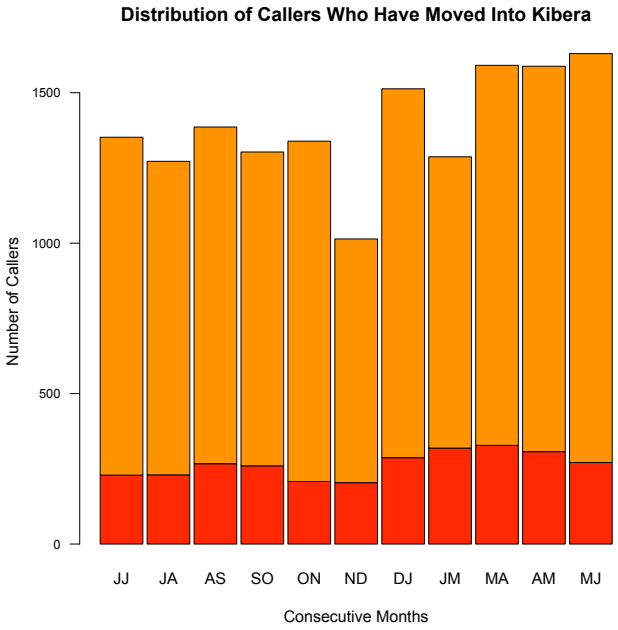

FIG. 16: Distribution of calls who have moved into Kibera. The orange bar represents those who have moved in from another part of Kibera. The red bar represents those who have moved into Kibera from outside of the slum.

We hope to better analyze movement within a slum. For example, do we see trends in movement around a slum? Does a person's place of work or tribal affiliation affect their movement to other villages in Kibera? These are question we are hoping to answer with future research.

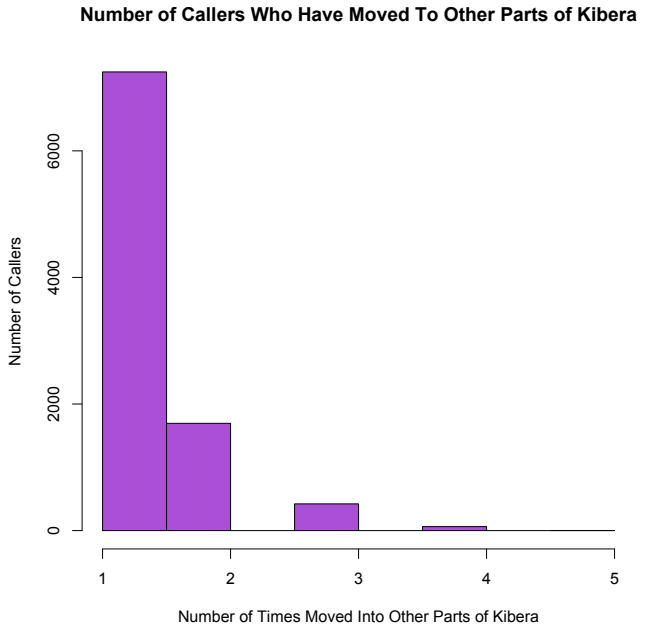

FIG. 17: This graph represents the number of times a person has moved to another part of Kibera and back to our first definition of living in Kibera. Here the mean is 1.29.

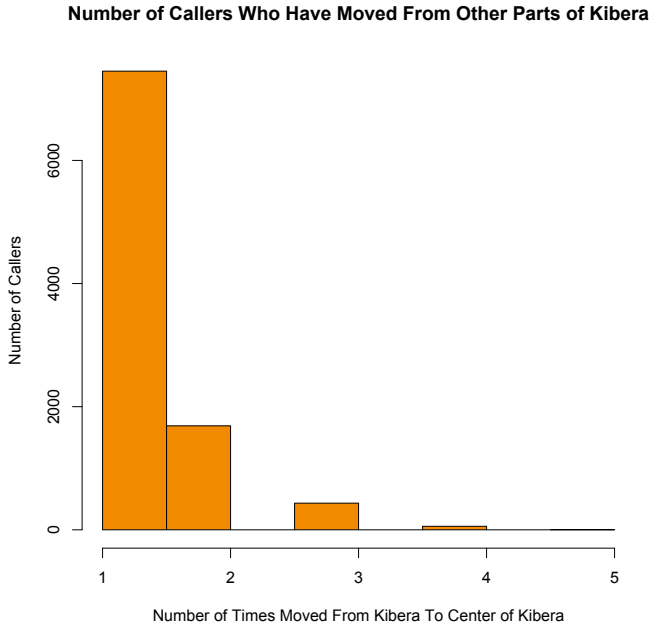

FIG. 18: This graph represents the number of times a person has moved from another part of Kenya and back to our first definition of living in Kibera. Here the mean is 1.284.

X. INFERRING PLACES OF WORK

Since work is the prime motivation for moving into Kibera, we wanted to better understand this driving factor. We are now considering all calls made to any tower besides those in T_e. Since the majority of these calls are made in Nairobi, we further refine our search to only those towers in Nairobi.

First, it is important to note that the percentages considered for each of these towers is extremely low with the maximum value at 0.008% of total calls. While these percentages are low, this fact is not surprising. Since the majority of calls made by slum residents are to others towers in the slum, shown in Figure 9, and the towers in Nairobi are extremely dense, we expect percentages to be low[8].

Our first method is to construct surface plots of Nairobi using the percentage of calls made at each site ID. With these surface plots, the main result is that there exists a strong connection between the center of Nairobi and Kibera. We have constructed surface plots, where the x coordinate is gridded latitude, the y coordinate is gridded longitude, and the z coordinate represents magnitute. We have set the center of Nairobi to be (0,0). In order to determine the x and y coordinates, we have subtracted the center of Nairobi's coordinates from each towers original longitude and latitude and multiplied the difference by 1000[9]. The z coordinates are the original percentage of calls multiplied by 1000[10]. From this, we have gridded all data onto a 150 by 150 surface. The magnitude of each square is then colored appropriately, with red being the most dense. Although we do not see overwhelming trends from month to month, the only consistent results we have found from this analysis is that the center of Nairobi has the densest number of towers. Below is an example of two plots generated using this method, in a flattened surface form.

Next we consider only the top 50 towers. We are looking at the overlap in these top towers with two different methods. The first method considers those towers with the greatest overlap in months. For example, we are focusing our analysis on towers that are listed in twelve, eleven, ten, and nine months out of the total number of months. For this analysis we first consider the distribution of the percentage of calls made at these towers over the course of the year. Finally, we plot the location of these towers in Google Earth to associate locations with possible working environments. These towers most likely are indicative of the most stable regions for work in Kibera.

First, let us consider calls made at towers that are listed in all twelve months. As we can see in Figure 21, the percentage of calls made in later months, i.e. April, May, and June 2009, are higher then the previous percentages. Most likely, this coincides with a higher number of calls made outside of the slum during these months. Also we have constructed similar graphs representing towers listed in nine, ten, and eleven months,

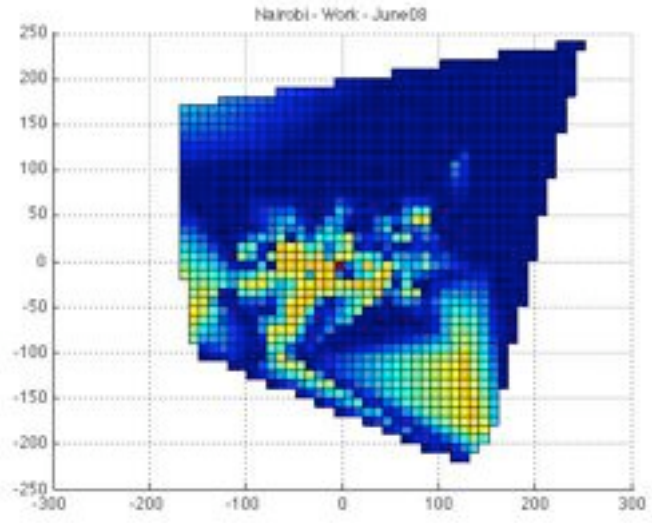

FIG. 19: This graph represents a Matlab plot of working regions in Nairobi for June 2008.

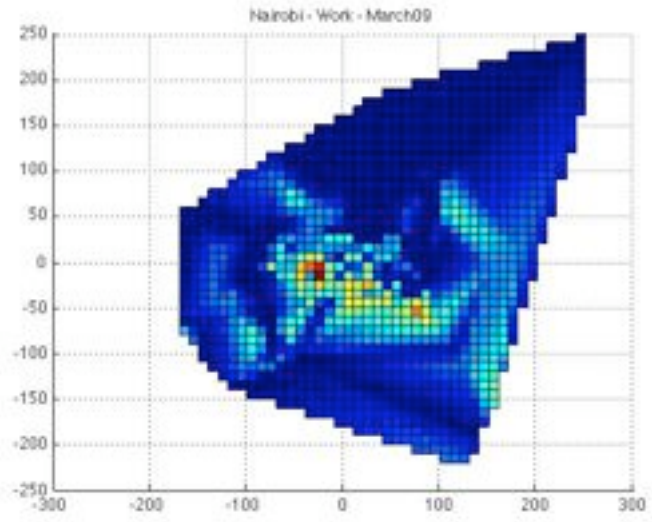

FIG. 20:

which are omitted from this paper. Overall, this method does not seem to given any insight into the prominent working regions of Kibera's residents.

On the other hand, this method enabled us to determine the average number of months a work region tower is listed. With the mean 1.950, we see that there is little overlap in work region towers from month to month.

The next method involves determining where these top work regions are located in proximity to Kibera. As Figure 22 shows, there appear to be two distinct groupings which are highlighted in Figures 23 and 24.

Our last method is to consider the top 5 towers called from for each month. Regardless of the number of overlap of these towers between months, we want to see where the

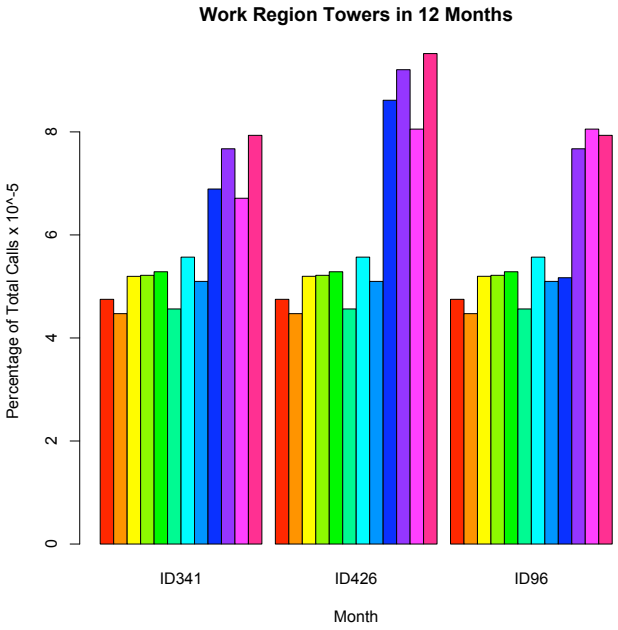

FIG. 21: This graph represents the three towers outside of Kibera listed in all twelve months. Each bar represents a the percentage of calls made at a given tower from month to month.

FIG. 22: This Google image highlights the top working towers occuring in the highest number of months. In the bottom left-hand corner is Kibera and the top right-hand corner is the center of Nairobi.

majority of people are going when they leave Kibera in each month. For this, we have plotted locations using Google Earth to get a broad idea of the the length of travel to each of these locations.

Overall, this method still shows a strong connection between Kibera and the center of Nairobi while allowing us to begin to analyze the distance traveled by slum residents. Although the distance shown in Figure 22 would suggest that Kibera's residents do not travel long distances for work, Figure 25 would tell a different story. We plan on using this result to help guide future research questions about the radius of mobility for slum residents.

FIG. 23: This Google image highlights one of the two major groupings of working towers. Each yellow placemarker is labeled according to the number of months this tower is listed. This grouping is in the center of Nairobi with many industrial places of work.

FIG. 24: This Google image highlights the second major grouping of working towers. Each yellow placemarker is labeled according to the number of months this tower is listed. This grouping is located southwest of the center of Nairobi with many industrial sites marked.

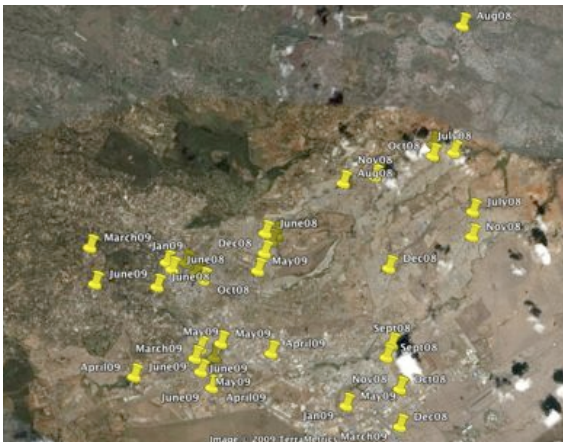

FIG. 25: This Google image shows the top 5 towers outside of the slum in each month. Each yellow placemarker is labeled according to the month the tower is listed.

FIG. 26: This Google image shows the first grouping of towers from Figure 25 of the center of Nairobi.

FIG. 27: This Google image shows the second grouping of towers from Figure 25 of an industrial region just south of the center of Nairobi.

FIG. 28: This Google image shows the third grouping of towers from Figure 25 to the southeast of the center of Nairobi.

FIG. 29: This Google image shows the fourth grouping of towers from Figure 25 to the east of the center of Nairobi.

XI. CONCLUSION

Although laws of human movement are difficult to harness, research using mobile phone tower locations and large sample sizes help address this problem. Out of all results, the most striking is the high turn over rate in Kibera. Using this methodology and data set, we found that Kibera has around 50% of individuals moving, whether to other parts of Kibera or elsewhere, per month. The transient nature of this population was evident in our analysis and has made harnessing any fundamental properties of slum residents difficult. When attempting to infer new places of residence, tribal affiliations, and work regions there seemed to be little overlap from month to month. We did find that a large proportion of working regions are in the center of Nairobi, which supports the theory that Kibera has been able to continue to grow as a result of its close proximity to the center of the city.

Our analysis has been conducted on a much larger scale than previous studies. Even though 18,000 people living in Kibera is lower than the estimated 600,000 residents, this sample is at least one hundred times larger than any previous study. Most importantly this work has raised new questions and avenues for future research.

XII. FUTURE WORK

There remains much work to be done. In this stage of research, the main goal was to get aggregate statistics to better understand the movement patterns of slum residents. We hope to use these statistics and intuition gained from this research to better understand the dynamics driving Kibera. Here is an overview of a number of projects we hope to work on.

A. Slum Dwellers Characteristics

First, we plan on determine overall statistics for all slum residents. We want to be able to know the total number of calls, length of calls, number of individual regions called, total number of tribal regions called at, distribution of calls over the day, and number of different people you're calling for all individuals classified as living in a slum. We are going to use this to compare to the average population of Kenya.

B. Migration

Associated with refining our definition of migration is a better assessment of tribal affiliation. Moreover, we want to be able to further classify those by:

- duration spent in the slum
- previous tribal region

Also, we want to be able to classify individuals moving into Nairobi in general, not necessarily slums. We want to be assess what percentage of migrants move into slums versus other parts of Nairobi.

- Determine the location of the first tower you can from in Nairobi
- Determine your top tower in Nairobi and its location
- If a person has made calls in Nairobi versus no calls ever made in Nairobi
- If you moved into Nairobi – only consider those people who have really moved there
- On average, how many months do you stay in Nairobi
- How mobile are you in Nairobi?
- Does your living region change over the year?
- Do you return to your home region? stay in Nairobi?

C. Change in Social Networks

There has been recent work on the influence of social exclusion in the growth of slums. Sociologists define social exclusion as: a multidimensional process of progressive social rupture, detaching groups and individuals from social relations and institutions, and preventing them from full participation in the normal, normatively prescribed activites of the society in which they live[13]. UN-HABITAT lists eliminating social exclusion of groups as one of the key steps to reducing the growth of slums. The intuition is that when individuals experience social exclusion, they are more likely to be lost among the masses and fall into poverty. Are the average slum residents more socially excluded than a random sampling of Kenyans? If an individual is experiencing more social exclusion, are they more likely to move into a slum?

Moreover, it is well accepted in sociology literature that when individuals move into an urban area their social network changes. Currently, Eagle et al. are working on quantifying this change using mobile phone data. However, we wonder if this change can also be seen in slums?

D. Analyzing Other Slums in Kenya

All of our work has focused on Kibera. However, we have identified other slums in Nairobi and plan on extending our methods to those places. From this, we hope to determine if there are general dynamics and rules governing these informal settlements across Nairobi. We

hope to be able to further extend this work to slums in other countries to see if trends found in Nairobi represent a universal rule for slums.

E. Stability of a Slum

With informal housing, temporary jobs, and constant migration, slums are inheriently unstable when compared to other residential areas. However, how can we develop a metric that will assess the stability of a slum? We have begun to address this question by looking at the influx of people in and out of Kibera. Moreover, we have begun to identify work regions and the stability of those regions. We hope to use this knowledge and other aggregate statistics from Kibera to develop a metric for the stability of Kibera and slums in general. We plan on using this metric to compare slums across Nairobi and other countries.

F. Categorization of Different Types of Slum Dwellers

Can we be able to pinpoint seasonal works, different demographics, those who work in a slum? Can we rank individuals by mobility? calling patterns? number of social contacts?

G. Social Network of a Slum

Is there an interconnected network of slums? Do slum residents tend to move into other slums? Is there a large amount of communication between slums in a city? Is there a large overlap of working regions? moving regions? How connected are individuals within a slum? Has a community formed? Are people living as mostly separate individuals?

H. Success of a Slum

Is a slum successful? Slums serve a specific function in a city. They provide temporary housing and a close proximity to work for individuals moving into urban areas. Slums can be considered successful if a large number of people flow through the area and are able to increase their socioeconomic status. We plan on using other mobile phone data, to assert if a person's socioeconomic status has increased during their time in the slum. Moreover, we plan on coupling this with the movement in and out of a slum to assert how successful a slum is and how its role is impacting the growth of the city?

I. Building a Generative Model

Using techniques from artificial intelligence, we plan on building a model that accurately reflects and predicts slum dynamics. In particular, we hope to be able to address specific questions of the growth of slums, i.e. where we think the next slum will form, how we think a slum will grow, and how the growth of a slum will affect the growth of the surrounding urban area.

J. Connection Between Original Region

How strong is the tie between an individual's region and their current residence? Is there a large flow of communication between a rural region and the slum?

[1] "Kenya - The Unseen Majority: Nairobi's Two Million Slum-Dwellers", Amnesty International, 2009.

[2] Brown, Lawrence A., Moore, Eric G.. "The Intra-Urban Migration Process: A Perspective". Geografiska Annaler. Series B, Human Geography, Vol 52, No. 1 (1970), pp. 1-13.

[3] F. Calabrese and C. Ratti. "Real Time Rome," Networks and Communication Studies - Official Journal of the IGU's Geography of Information Society Commission, 20:3, 247-258, 2006.

[4] Byerlee, Derek. *Rural-Urban Migration in Africa: Theory, Policy, and Research Implications*. International Migration Review, Vol. 8, No. 4 (Winter, 1974), pp. 543-566.

[5] Davis, Mike. *Planet of Slums*. London: Verso, 2002.

[6] http://www.gdrc.org/uem/squatters/define-squatter.html

[7] M. Gonzales, C. Hidalgo, and L. A. Barabasi, "Understanding Individual Human Mobility Patterns," Nature 453, pp. 779-782, 2008.

[8] Goodman, Allen E.. Reviewed work(s): *The Urbanization Process in the Third World: Explorations in Search of a Theory by T. C. McGee*. Economic Geography, Vol. 49, No. 1 (Jan., 1973), pp. 83-84.

[9] Hanna, William J., Hanna, Judith L.. *Urban Dynamics in Black Africa: An Interdisciplinary Approach*. Edison: Aldine Transaction, 2009

[10] Kramer, Mark. *Dispossessed: Life in Our World's Urban Slums.* Maryknoll: Orbis Books, 2006.

[11] Neuwirth, Robert. *Shadow Cities: A Billion Squatters, A New Urban World*. New York: Routledge, 2006.

[12] Nnadozie, Emmanuel U.. *African Economic Development*. New York: Academic Press, 2003.

[13] Silver, Hilary. "Social Exclusion: Comparative Analysis of Europe and Middle East Youth," Middle East Youth Initiative Working Paper (September 2007), p.15.

[14] Todaro, Michael P. "Urbanization, unemployment and migration in Africa: Theory and Policy," Policy Research Division Working Paper no. 104. New York: Population Council.

[15] Turner, John C. (1968) "Housing Priorities, Settlement Patterns, and Urban Development in Modernizing Countries", Journal of the American Planning Association, 34:6, 354-363

[16] UN Department of Economic and Social Affairs, Population Division, *World Urbanization Prospects*, the 2001 Revision, New York 2002.

[19] UN-HABITAT. *The Challenge of Slums - Global Report on Human Settlements*. London: Earthscan, 2003.

[18] UN-HABITAT. *The State of African Cities 2008*. New York: Earthscan, 2008.

[19] UN-HABITAT. *The State of the World's Cities 2008/2009*. London: Earthscan, 2009.

[20] Yue-man, Yeung, "Geograph in an Age of Mega-Cities," *International Social Sciences Journal* 151 (1997), p. 93.

[21] Van der Ploeg, F. and Poelhekke, S. "Globalization and the Rise of Mega-Cities in the Developing World," CESifo Working Paper No. 2208 (February 2008).

[1] http://www.brinq.com/projects/kenya/dialects.html

[2] We are missing February 2009 from our data set

[3] Note: all callers identifications are unique hashed IDs

[4] This constraint will eliminate outliers.

[5] We realize that our sample is not necessarily a random sampling of the entire population. One main objective is that we are only sampling individuals who can afford mobile phones. However, with the rate of mobile phone adoption and prevalence of cell phones in Kenya we believe that we our sampling is not entirely biased towards wealthier slum residents

[6] We have not fully analyzed C_2, although a preliminary analysis does not show any overwhelming evidence indicating where individuals are moving from. We plan on using new tools developed in our future work section to better analyze this set of individuals.

[7] These calls have occurred at towers not listed in our master tower files.

[8] We have also run our analysis by constructing regions (in the same fashion as the construction of living regions). However, we feel as though those results are less illustrative then the work C_3.

[9] We multiply by 1000 to reduce rounding errors.

[10] Once again, we multiply by 1000 to reduce rounding errors, since the original values are extremely low.